Authenticating Media Signals by Adjusting Frequency Characteristics to Reference Values

Related Application Data

This patent application claims the benefit of US Provisional Application 60/232,163 filed September 11, 2000, which is hereby incorporated by reference.

Technical Field

The invention relates to steganography, data hiding, and authentication of media signals, such as images and audio signals.

Background and Summary

Digital watermarking is a process for modifying physical or electronic media to embed a machine-readable code into the media. The media may be modified such that the embedded code is imperceptible or nearly imperceptible to the user, yet may be detected through an automated detection process. Most commonly, digital watermarking is applied to media signals such as images, audio signals, and video signals. However, it may also be applied to other types of media objects, including documents (e.g., through line, word or character shifting), software, multi-dimensional graphics models, and surface textures of objects.

Digital watermarking systems typically have two primary components: an encoder that embeds the watermark in a host media signal, and a decoder that detects and reads the embedded watermark from a signal suspected of containing a watermark (a suspect signal). The encoder embeds a watermark by altering the host media signal. The reading component analyzes a suspect signal to detect whether a watermark is present. In applications where the watermark encodes information, the reader extracts this information from the detected watermark.

Several particular watermarking techniques have been developed. The reader is presumed to be familiar with the literature in this field. Particular techniques for

embedding and detecting imperceptible watermarks in media signals are detailed in the assignee's co-pending application serial number 09/503,881 and US Patent 5,862,260, which are hereby incorporated by reference. Examples of other watermarking techniques are described in US Patent Application 09/404,292, which is hereby incorporated by reference. Additional features of watermarks relating to authentication of media signals and fragile watermarks are described in US Patent application 60/198,138, 09/498,223, 09/433,104, and 60/232,163, which is hereby incorporated by reference.

The invention provides a method of authenticating a media signal and related software, systems and applications. The method transforms at least a portion of the media signal into a set of frequency coefficients in a frequency domain. For example, it applies a Fast Fourier Transform (FFT) or other frequency transform to blocks of a media signal, such as an image, audio or video signal. It adjusts a relationship between selected frequency coefficients to a reference value. This adjustment is selected so that an alteration to be detected, such as a re-sampling operation or digital to analog- analog to digital conversion, alters the relationship. To detect the alteration, a detector computes the relationship in a potentially corrupted version of the signal. It then compares the result with a threshold value to detect whether the alteration has occurred.

A furtheraspect of the invention is a method of authenticating a media signal. The method evaluates signal peaks at selected frequency coefficients of the media signal. In a prior embedding process, the media signal has been modified to include peaks at the selected frequencies, such as by the technique summarized in the previous paragraph. The method determines, based on degradation of the signal peaks, whether the media signal has been altered. The frequency location of the peaks may vary from one application to the next. To detect, scanning and printing of watermarked images for example, the peaks are located at higher frequencies.

Another aspect of the invention is a watermark decoder, which includes a detector and analyzer for determining alteration of a watermarked media signal. The detector correlates a calibration signal with a media signal suspected of carrying a watermark to determine orientation parameters describing orientation of the media signal at embedding

of the watermark. The calibration signal includes a set of peaks at selected frequency coefficients. The analyzer orients the media signal using the orientation parameters and evaluates whether the media signal has been altered by examining signal peaks at selected frequency coefficients in the media signal.

Further features will become apparent with reference to the following detailed description and accompanying drawings. The following description details a method for detecting whether an image has been scanned, printed or photocopied after being processed by the method. It also describes alternative implementations and applications.

Brief Description of the Drawings

Fig. 1 is a flow diagram illustrating a process of embedding an authentication watermark in a media signal.

Fig. 2 is a flow diagram illustrating a process of detecting the authentication watermark from a potentially corrupted version of the watermarked signal.

Detailed Description

Fig. 1 is a flow diagram illustrating a process of embedding an authentication watermark in an input media signal (100), and in particular, in an image. The embedder begins by dividing a grayscale image into NxN blocks of samples at a specified resolution (102), where N is a pre-defined integer. For each block, the embedder computes a frequency transform of the image samples in that block (104), namely, a fast Fourier transform. From the mid-frequency and mid-high frequency coefficients, the embedder selects M Fourier transform coefficients (106), where M is a pre-defined integer. The coefficient locations are fixed by a pre-defined pattern. For example, the locations are scattered among roughly 25 to 100 coefficient locations in the mid to mid-high frequency range of a Fourier transform domain of a block of image samples where N ranges from 64 to 512 at spatial resolutions ranging from 75 to 600 dots per inch (DPI). The locations are symmetric about vertical and horizontal axes (and potentially diagonal axes) to facilitate detection as explained further below.

For each of the M selected coefficients, x, the embedder computes a ratio of the magnitude of a selected coefficient relative to the magnitude of its neighbors (108). In particular, it is a ratio of the magnitude of the selected coefficient to the average magnitude of the surrounding neighbors:

r(x) = Magnitude_of_x / Average_of_Magnitude_of_Eight_Neighbors_of_x

If r(x) < r, where r is a pre-defined reference value, the embedder increases the magnitude of x such that:

r(x) = r.

In this implementation, the value of r is a pre-defined constant. The reference may be derived dynamically from the input media signal. Also, the reference may be selected from a table of values so as to select the value of r in the table at the minimum distance from r(x). The adjustment to the host image is selected so as to be imperceptible or substantially imperceptible to a user in an output form of the watermarked signal.

Next, the embedder computes the inverse fast Fourier transform on each block to obtain the watermarked grayscale image (112). The watermarked image (114) may then undergo one or more transformations, such as digital to analog conversion, printing, scanning, analog to digital conversion, photocopying, etc. These transformations tend to corrupt the watermarked image in a predictable way.

The watermarking process of Fig. 1 may be combined with another watermarking process to embed other watermarks, either robust or fragile to transformations such as sampling distortions, geometric distortions, scaling, rotation, cropping, etc. In particular, the process may be combined with an embedding process described in pending application serial number 09/503,881 or US Patent 5,862,260 to encode a calibration signal that enables a detector to compensate for distortions such as scaling, rotation, translation, differential scale, shear, etc. In one implementation, for example, the calibration signal comprises an array of impulse or delta functions scattered in a pattern in the Fourier domain of each block of image samples. To embed the pattern, the embedder perceptually adapts the calibration signal to the host image block and adds it to that block. The impulse functions of the calibration signal have a pre-defined magnitude

and pseudo-random phase. To make the calibration signal less perceptible yet detectable, the embedder modulates the energy of the calibration signal according to the data hiding attributes (e.g., local contrast) of the image samples to which it is added. Preferably, the locations of the impulse functions are scattered across a range of frequencies to make them robust to transformations like spatial scaling, rotation, scanning, printing, and lossy compression. Further, they are preferably arranged to be symmetric about vertical and horizontal axes in the Fourier domain to facilitate detection after flipping or rotating the watermarked image.

The frequency coefficient locations selected for the method illustrated in Fig. 1 may be mutually exclusive or overlap the coefficient locations of the calibration signal. The calibration signal preferably has impulse functions at lower frequencies to survive compression, scanning, printing, etc. while the pattern of coefficients employed in Fig. 1 includes coefficients at locations that are likely to be impacted by alterations to be detected, such as printing, scanning and photocopying. In the case where they overlap, the modification of the coefficients according to Fig. 1 is implemented so as not to interfere with the calibrations signal. In particular, the embedder adjusts the selected coefficients as shown in Fig. 1 after the impulse functions of the calibration signal have been introduced, or the embedder calculates the watermarked signal taking into account the changes of the coefficient values due to the calibration signal and the process of Fig. 1.

Another approach is to adjust the selected frequency coefficients in the method of Fig. 1 so that those coefficients act as both a calibration signal and an authentication signal. The locations of the coefficients for the method of Fig. 1 and the delta functions of the calibration signal are the same. The embedder increases the magnitudes of selected mid to mid-high frequency coefficients relative to their neighbors to achieve the desired relationship with neighboring coefficients for authentication purposes. Since this modulation includes the addition of a delta function to the selected coefficients, it also inherently embeds a calibration signal comprised of delta functions at the selected locations. To compensate for rotation and scale, the detector performs a Fourier Mellin

transform of the suspect signal and the calibration signal into a log-polar space and then correlates the two signals. The location of the correlation peak in log polar space provides the spatial scale and rotation parameters. These parameters may then be used to compensate for rotation and scale changes before performing additional watermark decoding operations, such as the authentication operations of Fig. 2.

To compute translation, the delta functions added to the selected coefficients may be given a known pseudorandom phase. In this case, the detector correlates the phase information of the calibration signal with the suspect signal after compensating for rotation and scale. The location of the correlation peak gives the translation offset in the horizontal and vertical directions.

In addition to being integrated with other watermark signal components, the process of Fig. 1 may be combined with a robust watermark embedding process to carry a multi-bit message payload carrying metadata or a link to metadata stored in an external database. Example implementations for embedding this type of robust watermark are described in pending application serial number 09/503,881 and US Patent 5,862,260.

Fig. 2 is a flow diagram illustrating a process of detecting the authentication watermark from a potentially corrupted version of the watermarked media signal (120) from the process of Fig. 1. The first four steps (122) are the same as shown in the embedder. For each block, the detector computes the average of r(x), where x is over all M selected coefficients (124),

R = Average_of_r(x)

The detector computes the average of R over all blocks (126),

AR = Average_of_R

To detect whether the watermarked signal has undergone alterations, the detector compares the average of R with a pre-defined threshold (128). If AR >= T, where T is a pre-defined threshold, then the detector classifies it as original. If AR < T, then the detector classifies it as a copy. Depending on the application, the detector may indicate the result (130) to a user through some user interface (e.g., visual display, audio output such as text to speech synthesis, etc.). The detector may also indicate the result (130) to

another software process or device to take further action, such as communicating the event to a another device or database for logging, recording tracer data about the user or device in which the alteration is detected, linking the detecting device to a network resource such as a web site at a specified URL that informs the user about usage rules, licensing opportunities, etc.

To make the process robust to geometric distortion, the detector includes a pre-processing phase in which it correlates a calibration signal with the potentially corrupted watermarked signal as described in pending application serial number 09/503,881 or US Patent 5,862,260. Using a Fourier Mellin transform, the detector maps both the calibration signal and the received signal into a log polar coordinate space and correlates the signals (e.g., using generalized matched filters) to calculate estimates of rotation and scale. After compensating for rotation and scale, the detector uses the phase information of the calibration signal to compute translation, e.g., the origin or reference point for each block. Further correlation operations may be used to compute differential scale (e.g., the change in scale in the horizontal and vertical directions after watermarking). After compensating for geometric distortion, the detector executes the process of Fig. 2 to detect alteration in the selected frequency coefficients modified according to the method shown in Fig. 1.

While the invention is illustrated with respect to a specific implementation, it may be implemented in a variety of alternative ways. For example, the above example specifically refers to a grayscale image. This example may be adapted to other types of images including video and still imagery, color and monochrome images, etc. For color images, the embedding and detecting operations may be performed on two or more color channels, including luminance, chrominance or some other color channels. The embedding and detecting operations may be applied to frequency coefficients of alternative frequency transforms, such as DCT and wavelet, to name a few.

The embedding process shown in Fig. 1 may be performed on a portion of the host signal to create a watermark signal that is combined with the host signal. For example, in one possible implementation, the embedder pre-filters the host signal to yield

a high pass filtered signal including content at the mid and high frequency ranges impacted by the watermark. The embedder makes the modification to this filtered signal, and then combines the resulting modified signal with the original signal.

The embedding and detecting processes may also be integrated into compression and decompression operations. For example, the frequency domain transform may be executed as part of a compression process, such as JPEG, JPEG 2000 or MPEG, where blocks of the signal are transformed into a frequency domain. Once converted to the frequency domain, frequency coefficients may be adjusted as described above.

The embedding and detecting operations apply to other media types, including audio media signals. In addition, the frequency domain coefficients may be selected and adjusted to reference values to detect other types of signal alteration, such as lossy compression, digital to analog and analog to digital conversion, downsampling and upsampling, etc.

Semi-fragile watermarks

A related watermarking approach is to use an array of Fourier magnitude impulse functions with random phase (a calibration signal, also referred to as a watermark synchronization or orientation signal) for semi-fragile, and copy and copy-attack resistant watermarks. Semi-fragile refers to a watermark that degrades in response to some types of degradation of the watermarked signal but not others. In particular for document authentication applications using such a watermark, the watermark decoder can determine if the watermark has been scanned and printed or battered by normal usage, potentially while being read with a web camera. The copy-attack relates to the assertion that one can use noise-reduction, i.e. Wiener filters, to lift a watermark and, then using threshold and masking techniques, one can re-embed it in a new image. Interestingly, these concepts are related because they both include an additional scanning and printing cycle, assuming the copy attack works on printed, not only digital, content. This type of semi-fragile watermark can be used to determine if a watermarked document has been copied, possibly using a high quality copier and low quality reader, and as such, can stop copying and can be used to measure quality of service.

One approach to implementing a semi-fragile watermark is to embed extra signal peaks in the Fourier magnitude domain that are of varying intensity, and have the watermark decoder determine if the watermark has been scanned and printed by the relative power of the extra and original calibration signal peaks. The extra peaks refer to a set of peaks used to implement the semifragile watermark. The original calibration signal peaks refer to the ones already included in the watermark to determine its orientation in a geometrically distorted version of the watermarked signal. For an example of such a calibration signal, see U.S. Patent No. 5,862,260 and U.S. Application No. 09/503,881, which are incorporated by reference. Some peaks are referred to as "extra" because they are included in addition to other peaks that form the original set of peaks in a calibration signal.

One advantage of including the semi-fragile watermark in the calibration signal is that the robust part of the watermark, which includes the desired detailed information such as a unique ID, is the multi-bit message, whereas the fragile part, which is only used to determine copying, is a few bit message. The fragile watermark can be considered as a single bit (copied or not) but actually allows more information by being frequency specific, as described below. Interestingly and potentially advantageously, the semi-fragile watermark is separate but inherently related to the robust watermark -- thus they cannot be separated for successful copy attacks.

Specifically, the extra calibration signal peaks should be located at frequencies that best discriminates between the printing and scanning process, normal scuffing and a web camera reader. These locations can be determined by analyzing the frequency response of printing, scanning, scuffing and web cameras for frequency differences.

For example, a printing-scanning process may represent high-frequencies better than a camera, but not low frequencies. In addition, scuffing may show low-and-high frequency losses. Thus, the reader will be able to determine if the watermark has been copied, involving an additional scanning-printing process, by the relative intensities of the extra and original calibration signal peaks at low and high frequencies. In this example, high-and-low frequency loss is acceptable, whereas only low frequency loss

represents a copied watermark.

In addition, the extra calibration signal peaks could also be dependent upon the content of the host signal, thus providing additional defense against the copy attack. For example, the host image samples could be broken in 16 equal sub-blocks, and the location of the extra peaks depends upon the average intensity of each quadrant to the total average intensity. Or, if only a section of the image is visible to the reader, each 32 by 32 sample block could be used in the above calculation instead of the complete image. Any "hash" of the host image that survives a web camera reader (referred to as a perceptual hash) could be used. To this end, if the watermark is moved to another picture, after it is read, it is less likely that the extra calibration signal peak locations are correct, not to mention that the less intense calibration signal points have been removed by the additional scanning-printing process.

Alternatively, in regards to the copy attack, the content dependent information could be used to slightly move the location of a few original calibration signal peaks, as opposed to adding extra calibration signal peaks. This means that the image content is implicitly in the calibration signal's jitter, and the copy attack is less likely to succeed unless the read and embedded images have the same perceptual hash. On the one hand, this approach may reduce robustness of the robust message to scaling, rotation and translation. On the other hand, no extra bits containing the output of the perceptual hash need to be embedded in the robust message.

Concluding Remarks

Having described and illustrated the principles of the technology with reference to specific implementations, it will be recognized that the technology can be implemented in many other, different, forms. To provide a comprehensive disclosure without unduly lengthening the specification, applicants incorporate by reference the patents and patent applications referenced above.

The methods, processes, and systems described above may be implemented in hardware, software or a combination of hardware and software. For example, the

embedding processes may be implemented in a programmable computer or a special purpose digital circuit. Similarly, detecting processes may be implemented in software, firmware, hardware, or combinations of software, firmware and hardware. The methods and processes described above may be implemented in programs executed from a system's memory (a computer readable medium, such as an electronic, optical or magnetic storage device).

The particular combinations of elements and features in the above-detailed embodiments are exemplary only; the interchanging and substitution of these teachings with other teachings in this and the incorporated-by-reference patents/applications are also contemplated.